Code No: M155AC **R18**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech III Year I Semester Examinations, January/February - 2023
INTRODUCTION TO DATA SCIENCE
(Minor Program in Data Science)

Time: 3 Hours **Max. Marks: 75**

Note: i) Question paper consists of Part A, Part B.
ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.
iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A

(25 Marks)

1.a)	What is R Software?	[2]
b)	What is the difference between database and big data?	[3]
c)	Give an example for nominal attribute.	[2]
d)	What is the inter-quartile range and how is it calculated?	[3]
e)	Define an array.	[2]
f)	What are ordered factors in R?	[3]
g)	Give the syntax of 'for' loop and 'while' loop in R.	[2]
h)	Present a short note on nested functions in R.	[3]
i)	Write a short note on a regression model.	[2]
j)	What is the need of attribute subset selection technique?	[3]

PART – B

(50 Marks)

2.a) Compare population with sample.
b) Describe the concept of datafication. [5+5]

OR

3. How to create an R environment and run RStudio? Explain. [10]

4.a) Give a brief summary on graphic display of basic statistical description of data.
b) Differentiate between discrete and continuous attributes. [6+4]

OR

5. Classify the various common measures of central tendency and briefly explain them. [10]

6.a) Explain how to convert lists to vectors.
b) Examine if a factor is an ordered factor or not in R programming with an example. [5+5]

OR

7.a) Explain how to sort dataframes in R programming.
b) Describe how to create and name the vectors. [5+5]

8. Explain the following terms in R:
   a) Relational operators and vectors.
   b) Conditional statement. [5+5]

OR

9.a) Write R code for sum of series using recursion.
   b) Describe looping statements in R. [5+5]

10.a) Write short notes on data cube aggregation.
   b) Describe the icon-based visualization techniques. [5+5]

OR

11. Explain the following terms:
   a) Principal components analysis
   b) Visualizing complex data. [5+5]

---ooOoo---